

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

9/914870  
9/04/01

EP 00/01853

# BUNDESREPUBLIK DEUTSCHLAND

09/914870

4



REC'D 19 JUL 2000

WIPO

PCT

*F 900 / 1853*

## Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

**Aktenzeichen:** 199 58 979.8

**Anmeldetag:** 08. Dezember 1999

**Anmelder/Inhaber:** Marcus Hartmann,  
Münster, Westf/DE

**Bezeichnung:** Saure Hydrolasen sowie diese kodierende  
DNA-Sequenz aus Ciliaten und deren  
Verwendung

**IPC:** C 12 N 9/14

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Anmeldung.

München, den 15. Juni 2000  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

*fuß*

Seller



## **Saure Hydrolasen sowie diese kodierende DNA-Sequenz aus Ciliaten und deren Verwendung**

Die Erfindung betrifft eine Nukleinsäure kodierend für eine saure Hydrolase und die Verwendung dieser entsprechend dem Oberbegriff des Anspruchs 1 bis 19.

Die Expression von Fremdproteinen in Mikroorganismen, wie Bakterien, Hefen, oder Mammalia-Zellen ist für die biotechnologische Darstellung und Produktion rekombinanter Proteine von großer Bedeutung. Hierbei werden bakterielle Expressionssysteme auf der Basis von *E. coli* oder *B. subtilis* zur Produktion von rekombinanten Peptiden bzw. Proteinen wie z.B. Insulin, Interleukin-2, Gewebe-Plasminogen-Aktivator, Proteasen und Lipasen verwendet. Bei gram-negativen Bakterien basieren die Expressionssysteme meist auf der Verwendung von genetischen Elementen, wie dem lac-Operon oder dem Tryptophan-Operon. Dabei werden die wirtsfremden Proteine entweder in "inclusion bodies" (Einschlusskörper) innerhalb der Zelle produziert, oder, bei Verwendung von Expressionssystemen auf der Basis von  $\beta$ -Lactamase-Genen, in den periplasmatischen Raum. Die Produktion von rekombinanten Proteinen in das umgebende Fermentationsmedium ist nicht etabliert. Bei gram-positiven Bakterien werden bisher fast ausschließlich nur zelleigene Proteine in Expressionssysteme eingebaut und expremiert.

Hefen, wie *S. cerevisiae*, *Kluyveromyces lactis* oder *Pichia pastoris* werden ebenfalls zur heterologen Expression rekombinanter Proteine, wie z.B. von humanem Faktor XIIIa, bovinem Pro-Chymosin oder Oberflächenantigenen eingesetzt. Die Expressionssysteme basieren hier auf Shuttle-Vektoren (Vektor mit einem Hefe- und einem bakteriellen Anteil), die auf den genetischen Elementen der Galakto-Kinase-Epimerase,

der sauren Phosphatase oder Alkoholdehydrogenase aufgebaut sind. In der Regel wird das rekombinante Protein in das Cytoplasma der Zelle produziert. Bei Verwendung von hefeeigenen Signalsequenzen, wie der des Alpha-Faktors, können die exprimierten Proteine auch in das Fermentationsmedium sezerniert werden. Die Glykosylierung sezernierte Proteine erfolgt nach dem "high-mannose"-Typ.

Mammalia-Zellen, wie verschiedene Zelltypen von Nagern (CHO-Zellen, C127-Zellen) oder Affen (Vero-, CV-1- oder COS-Zellen) kommen ebenfalls zur heterologen Expression rekombinanter Proteine zum Einsatz. Die Expressionssysteme basieren hier auf rekombinanten Viren (BPV-Vektor) oder auf Shuttle-Vektoren. Zur Regulation der Expression werden virale SV40-Enhancer/Promotor-Systeme oder zelluläre Verstärker-Elemente eingesetzt. Die rekombinanten Proteine, wie Erythropoietin, werden in das Fermentationsmedium sezerniert, da die die Fremd-gene in der Regel bereits Signalsequenzen mitbringen und diese vom Expressionssystem verstanden und zur Zielsteuerung verwendet werden.

Zur biotechnologischen Produktion von glykosylierten, extrazellulären Enzymen werden desweiteren Ciliaten wie *Tetrahymena* eingesetzt. Ciliaten wachsen auf kostengünstigen Fermentationsmedien unter Einsatz von Standardfermentationsverfahren. Zur Transformation solcher Ciliaten stehen Vektoren zur Verfügung, die auf den rDNA-Elementen des Ciliaten *Tetrahymena* basieren. Zur heterologen Expression von bakteriellen Proteinen in Ciliaten werden DNA-Konstrukte aus Genen von *Tetrahymena* eingesetzt. Wenn geeignete genetische Elemente zur Regulation der Transkription, der Zielsteuerung und der Glykosylierung von Fremdproteinen zur Verfügung stehen, sind Ciliaten ein ideales Expressionssystem zur kostengünstigen Produktion therapeutischer re-

kombinanter Proteine.

Die bisher verwendeten Gram negativen bakteriellen Expressionssysteme führen in der Regel zur Bildung von "inclusion bodies" in der Zelle, einhergehend mit einer Denaturierung der Proteine. Für die Gewinnung des rekombinanten Proteins müssen die Zellen aufgebrochen und das denaturierte, inaktive Protein für seine Funktion zurückgefaltet werden. Dies verursacht zusätzliche kostenintensive Verfahrensschritte und erniedrigt die Ausbeute des erwünschten Proteins. Die für eukaryontische Proteine wichtige Glykosylierung unterbleibt völlig. Bei der Verwendung von Gram-positiven bakteriellen Expressionssystemen ist der Abbau des Zielproteins durch hohe proteolytische Aktivitäten in der Fermentationsbrühe zudem problematisch.

Bei der Verwendung von Hefen zur heterologen Expression wird das erwünschte Zielprotein häufig nur in die Zelle produziert, aus der es durch Zellaufschluß entfernt werden muss. Dies verursacht, wie bei bakteriellen Expressionssystemen, zusätzliche zeit- und kostenintensive Verfahrensschritte. Werden hefeeigene Signalpeptide verwendet, so werden die Fremdproteine bei der Sezernierung nicht korrekt gespleißt und glykosyliert.

Werden hingegen Mammalia-Zellsysteme zur Produktion von rekombinanten Proteinen eingesetzt, so liegen die erwünschten Proteine extrazellulär, korrekt gespleißt und glykosyliert im Fermentationsmedium vor. Nachteilig ist hier jedoch zum einen die niedrige Expressionsrate durch die fehlerhafte Prozessierung und ineffiziente Translation von Genen, die über virale Vektoren in das Genom der Produktionszelllinie eingebracht wurden. Zum anderen sind die serumhaltigen Fermentationsmedien für Mammalia-Zellen extrem kostenintensiv. Zudem ist die Fermentationstechnologie für die scherkraftempfindlichen Zelllinien auf

Grund von Konstruktionen zur blasenfreien Belüftung aufwendig und ebenfalls kostspielig. Weitere Probleme ergeben sich durch die hohe Gefahr der Infektion der Zelllinien durch Mycoplasmen und Viren. Zusammengekommen hat die Verwendung von Mammalia-Zellen zur biotechnologischen Herstellung von rekombinanten Proteinen sehr hohe Kosten, Sicherheitsauflagen und niedrige Ausbeuten zur Folge.

Für den Einsatz von Ciliaten, wie *Tetrahymena*, gelten die oben genannten Nachteile bei der Produktion von Proteinen nicht. So werden beispielsweise saure Hydrolasen, die an der Verdauung von Nahrungspartikeln beteiligt sind, in großen Mengen und komplex glykosiliert aus der Zelle ausgeschleust.

Bisher ist es jedoch nicht möglich, andere, fremde glykosylierte eukaryontische Proteine in Ciliaten zur Expression zu bringen, die gleichfalls in das Fermentationsmedium sezerniert werden. Ursache ist, dass bisher die DNA-Sequenzen aus Ciliaten eigenen sauren Hydrolasen unbekannt waren, die für die Konstruktion von Expressionsvektoren notwendig sind.

Aufgabe der Erfindung ist es, eine DNA-Sequenz von sauren Hydrolasen von Ciliaten abzugeben. Die DNA-Sequenz soll es ermöglichen, in einem Expressionssystem heterologe Proteine nach der Transformation in Ciliaten in das Fermentationsmedium zu verbringen. Dieses System soll auch eine hohe Expressionsrate des heterologen Proteins erreichen, das unter Kulturbedingungen in großen Mengen aus der Zelle ausgeschleust wird.

Diese Aufgabe wird durch ein System gelöst, bei dem eine Nucleinsäure mit einer Sequenz mit der Seq. ID. Nr. 1 codierend für saure Hydrolasen Verwendung findet. Erfindungsgemäß ist dies insbesondere eine

## DNA-Sequenz von sauren Hydrolasen

Die erfindungsgemäße DNA-Sequenz von sauren Hydrolasen beinhaltet insbesondere ein Signal- und ein Propeptid, sowie gegebenenfalls weitere genetische Elemente zur Zielsteuerung von Proteinen. Die Verwendung dieser Sequenzen in einem Vektor ermöglicht es, heterolog exprimierte Proteine aus der Zelle zu transportieren und damit ohne Zellaufschluß aus Fermentationsbrühe aufzureinigen.

Erfindungsgemäß wird insbesondere eine Nukleinsäure kodierend für eine saure Hydrolase aus Ciliaten (EC 3.-.-) beansprucht. Vorzugsweise ist dies eine Nukleinsäure kodierend für eine extrazelluläre saure Hydrolase aus Ciliaten, z. B. eine Nukleinsäure kodierend für eine saure Hydrolase, die eine Esterase ist (EC 3.1.-.-). Als Alternative kommen erfindungsgemäß auch eine Nukleinsäure kodierend für eine saure Hydrolase, die eine Glykosidase ist (EC 3.2.-.-) eine Nukleinsäure kodierend für eine saure Hydrolase, die eine Ether-Hydrolase ist (EC 3.3.-.-) eine Nukleinsäure kodierend für eine saure Hydrolase, die eine Peptid-Hydrolase ist (EC 3.4.-.-) eine Nukleinsäure kodierend für eine saure Hydrolase, die Kohlenstoff-Stickstoff-Bindungen hydrosyliert, die nicht peptidische Bindungen sind (EC 3.5.-.-). Desweiteren können erfindungsgemäß Nukleinsäuren kodierend für eine saure Hydrolase, die saure Anhydride hydrosyliert (EC 3.6.-.-), für eine saure Hydrolase, die Kohlenstoff-Kohlenstoff-Bindungen hydrosyliert (EC 3.7.-.-) für eine saure Hydrolase, die Kohlenstoff-halide Bindungen hydrosyliert (EC 3.8.-.-) eingesetzt werden.

In einer weiteren Ausführungsform der Erfindung kommen Nukleinsäuren kodierend für eine saure Hydrolase, die Phosphor-Stickstoff-Bindungen hydrosyliert (EC 3.9.-.-), Nukleinsäuren kodierend für eine saure Hydrolase, die Schwefel-Stickstoff-Bindungen hydrosyliert (EC



3.10.-.-), Nukleinsäure kodierend für eine saure Hydrolase, die Kohlenstoff-Phosphor-Bindungen hydrosyliert (EC 3.11.-.-) sowie Nukleinsäuren kodierend für eine saure Hydrolase, die Schwefel-Schwefel-Bindungen hydrosyliert (EC 3.12.-.-).

Figur 1 zeigt eine Nukleinsäure kodierend für eine saure Hydrolase aus Ciliaten. Figur 2 zeigt ein entsprechendes Expressionsprodukt der Nukleinsäure gemäß Seq. ID. Nr. 1. Dieses Protein gemäß Seq. ID. Nr. 2 ist ebenfalls Gegenstand der vorliegenden Erfindung.

Die Verwendung einer Nucleinsäuresequenz von sauren Hydrolasen gemäß der Erfindung oder Teilen davon zur homologen oder heterologen Expression von rekombinanten Proteinen und Peptiden, sowie zur homologen oder heterologen Rekombination ("knock-out", "gene replacement").

Gegenstand der Erfindung ist ein Verfahren, bei dem die für saure Hydrolasen kodierende erfindungsgemäße Nukleinsäure oder Teile davon mit den für homologe oder heterologe Expression üblichen Enhancer, Promotoren, Operatoren, Origins, Terminatoren, Antibiotikaresistenzen oder anderen Nukleinsäuren bzw. DNA-Fragmenten bzw. Sequenzen jeglicher Art von Viroiden, Viren, Bakterien, Archezoen, Protozoen, Pilzen, Pflanzen, Tieren oder Menschen kombiniert werden.

Insbesondere wird die erfindungsgemäße Nucleinsäure oder Teile davon in einen Vektor, ein Plasmid, ein Cosmid, ein Chromosom oder Minichromosom, ein Transposon, ein IS-Element, eine rDNA oder jede andere Art ringförmiger oder linearer DNA oder RNA eingebaut.

Dem Fachmann ist klar, dass erfindungsgemäß auch Nucleinsäuren eingesetzt werden können, die zu mindestens 40% homolog mit der Nuc-

leinsäure gemäß Seq. ID. Nr. 1 ist. Auch das Protein gemäß Seq. ID. Nr. 2 kann modifiziert werden ohne die Funktion zu verlieren. So können beispielsweise sogenannte konservative Austausche von Aminosäuren durchgeführt werden. Dazu können z.B. hydrophobe Aminosäuren untereinander ausgetauscht werden.

Zur Aufreinigung und Isolierung von sauren Hydrolasen aus Ciliaten und zur Bestimmung der Sequenz können folgende Methoden verwendet werden. Dies wird an den folgenden Beispielen verdeutlicht.

#### Beispiel 1

Aus insgesamt 3,2 L Zellkultur wurden Zellen des Ciliaten *Tetrahymena* der spätlogarithmischen Wachstumsphase in 400 mL Hungermedium (10 mM TRIS-HCl, pH 7,4) hineingewaschen und die Zellen unter Schütteln weitere 4 Stunden inkubiert. Anschließend wurde der zellfreie Kulturüberstand geerntet und durch eine Reihe Filter mit abnehmendem Porendurchmesser filtriert, um eventuell vorhandenes partikuläres Material zu entfernen. Das Filtrat wurde mit einer Amicon-Ultrafiltrationszelle eingengt und in den Startpuffer für die anschließende Ionenaustausch-Chromatographie (IEC) umgepuffert.

Die aufgefangenen Chromatographie-Fractionen wurden mit spezifischen 4-Nitro-phenyl-Substraten für saure Hydrolasen, der sauren Phosphatase (sPase) und der  $\beta$ -Hexosaminidase ( $\beta$ -Hex), getestet und die Fractionen mit der höchsten  $\beta$ -Hex-Aktivität vereinigt. Diese wurden durch eine weitere Amicon-Ultrafiltration eingengt und in den Startpuffer für die Affinitäts-Chromatographie umgepuffert. Die aufgefangenen Chromatographie-Fractionen wurden, wie oben beschrieben, auf Enzymaktivitäten getestet, die Fractionen mit der höchsten Aktivität für eine saure Hydrolase vereinigt und in Phosphat-Puffer (PB) umgepuffert.

Aus insgesamt acht Aufreinigungen wurden die hydrolase-haltigen Chromatographie-Fractionen zusammengefasst, eingengt, portioniert und bis zur weiteren Charakterisierung eingefroren.

Ein Teil dieser so aufgereinigten Hydrolase wurde über 2D-SDS-Gelelektrophorese aufgetrennt (insgesamt acht Gele), der Hauptspot jeweils ausgestochen und gesammelt. Das in diesen Gelstücken vorhandene Protein, die  $\beta$ -Hex, wurde mit der Protease Trypsin verdaut. Die so entstandenen Peptid-Fragmente wurden aus den Gelstückchen herausgelöst und über Reverse-Phase-HPLC (RP-HPLC) voneinander getrennt. Die Reinheit der HPLC-Fractionen wurde mit einem massenspektroskopischen Verfahren (MALDI-MS) überprüft und die sauberen, d.h. die Fractionen, die nur eine Peptid-Spezies einer definierten Masse enthielten, über Edman-Abbau von ihrem N-Terminus her sequenziert.

Anhand der Sequenz der Peptidfragmente, die durch den Trypsin-Verdau gewonnen wurden, wurden  $\beta$ -Hexosaminidase-spezifische PCR-Primer erstellt. Durch Sequenz-Vergleich mit den bereits vorhandenen  $\beta$ -Hex-Sequenzen konnte eine Vorauswahl der Primer-Kombinationen für die RT-PCR getroffen werden. Mit diesen Informationen wurde mittels RT-PCR aus isolierter Gesamt-RNA, deren Qualität vorher durch Northern-Hybridisierung überprüft wurde, erste spezifische cDNA-Fragmente amplifiziert und sequenziert. Mit Hilfe dieser Fragmente wurden sequenzspezifische Primer konstruiert, die für weitere PCR-Experimente eingesetzt wurden. Jeder Primer und jede gewonnene Teilsequenz wurde durch Datenbank-Abgleich und vor weiteren Experimenten u.a. auf eventuell übersehene Vektorsequenzen überprüft. Durch 5'- und 3'-RACE wurde die cDNA-Sequenz zum 5'- und 3'-Ende verlängert. Die so vervollständigte Sequenz der  $\beta$ -Hex-cDNA diente dann als Basis für weitere Sequenzanalysen.

Die so ermittelte Sequenz einer sauren Hydrolase von aus dem Ciliaten lautet, wie in Figuren 1 aufgeführt. Bei der Sequenz handelt es sich um insgesamt 1836 Basenpaare inkl. 5'- und 3'-nichttranslatierten Bereichen der  $\beta$ -Hexosaminidase mit einem offenen Leseraster ("open reading frame") von 1656 Basenpaaren Länge. Die komplementäre Aminosäuresequenz ist 551 Aminosäuren lang und lautet, wie in Figuren 2 aufgeführt.

Das Sequenz der sauren Hydrolase besitzt insgesamt 9-Glykosylierungsstellen und enthält ein Signalpeptid, sowie eine pro-Sequenz zur Zielsteuerung des Enzyms durch den Sortiermechanismus der Zelle.

### **Patentansprüche**

1. Eine Nukleinsäure kodierend für eine saure Hydrolase aus Ciliaten (EC 3 - - -).
2. Eine Nukleinsäure kodierend für eine extrazelluläre saure Hydrolase aus Ciliaten.
3. Eine Nukleinsäure kodierend für eine saure Hydrolase, die eine Esterase ist (EC 3.1.-.-).
4. Eine Nukleinsäure kodierend für eine saure Hydrolase, die eine Glykosidase ist (EC 3.2.-.-).
5. Eine Nukleinsäure kodierend für eine saure Hydrolase, die eine Ether-Hydrolase ist (EC 3.3.-.-).
6. Eine Nukleinsäure kodierend für eine saure Hydrolase, die eine Peptid-Hydrolase ist (EC 3.4.-.-).
7. Eine Nukleinsäure kodierend für eine saure Hydrolase, die Kohlenstoff-Stickstoff-Bindungen hydrosyliert, die nicht peptidische Bindungen sind (EC 3.5.-.-).
8. Eine Nukleinsäure kodierend für eine saure Hydrolase, die saure Anhydride hydrosyliert (EC 3.6.-.-).
9. Eine Nukleinsäure kodierend für eine saure Hydrolase, die Kohlenstoff-Kohlenstoff-Bindungen hydrosyliert (EC 3.7.-.-).
10. Eine Nukleinsäure kodierend für eine saure Hydrolase, die Kohlen-

stoff-halide Bindungen hydrosyliert (EC 3.8.-.-).

11. Eine Nukleinsäure kodierend für eine saure Hydrolase, die Phosphor-Stickstoff-Bindungen hydrosyliert (EC 3.9.-.-).
12. Eine Nukleinsäure kodierend für eine saure Hydrolase, die Schwefel-Stickstoff-Bindungen hydrosyliert (EC 3.10.-.-).
13. Eine Nukleinsäure kodierend für eine saure Hydrolase, die Kohlenstoff-Phosphor-Bindungen hydrosyliert (EC 3.11.-.-).
14. Eine Nukleinsäure kodierend für eine saure Hydrolase, die Schwefel-Schwefel-Bindungen hydrosyliert (EC 3.12.-.-).
15. Eine Nukleinsäure kodierend für eine saure Hydrolase aus Ciliaten, gekennzeichnet durch die Nukleinsäuresequenz in Fig. 1 oder Teile davon bzw. durch die Aminosäuresequenz in Fig. 2 oder Teile davon.
16. Ein Protein mit der Aminosäuresequenz der in Figuren 2 angegebenen Abfolge von Aminosäuren.
17. Verwendung einer DNA-Sequenz von sauren Hydrolasen gemäß den Ansprüchen 1 bis 16 oder Teile davon zur homologen oder heterologen Expression von rekombinanten Proteinen und Peptiden, sowie zur homologen oder heterologen Rekombination ("knock-out", "gene replacement").
18. Ein Verfahren, bei dem die Nukleinsäure von sauren Hydrolasen aus in den Ansprüchen 1 bis 16 oder Teile davon mit den für homologe oder heterologe Expression üblichen Enhancer, Promotoren,

Operatoren, Origins, Terminatoren, Antibiotikaresistenzen oder anderen Nukleinsäuren bzw. DNA-Fragmenten bzw. Sequenzen jeglicher Art von Viroiden, Viren, Bakterien, Archezoen, Protozoen, Pilzen, Pflanzen, Tieren oder Menschen kombiniert wird.

19. Ein Verfahren, bei dem eine Nukleinsäure von sauren Hydrolasen in den Ansprüchen 1 bis 16 in einen Vektor, ein Plasmid, ein Cosmid, ein Chromosom oder Minichromosom, ein Transposon, ein IS-Element, eine rDNA oder jede andere Art ringförmiger oder linearer DNA oder RNA eingebaut oder verwendet wird.

### **Zusammenfassung**

Eine Nukleinsäure kodierend für eine saure Hydrolase aus Ciliaten (EC 3 - - -).





GAA GAA TAA TGC TGG AAT AAA CGC CCT GAA ATT AAG GAA TTC ATG AAT TAA AAT AAC ATC  
1080

TCT ACA TAT ACT GAT TTG TAG AAT TAT TAC AGA AAG AAC TAA GTT AAC ATT TGG AAA TCA  
1140

ATT AAT GCT ACT AAG CCT GCT ATT TTC TGG GCA GAT TCA AAT ACT TTG AAA TAT GGT CCT  
1200

GAT GAT ATT ATT CAA TGG TGG GGA TCT ACT CAT GAT TTT TCT TCA ATC AAA GAT CTT CCT  
1260

AAC AAA ATA ATT TTA TCT TTC TAT GAT AAT ACT TAT TTG GAT GTT GGT GAG GGA AAT AGA  
1320

TAT GGT GGA AGT TAT GGC AGC ATG TAT AAC TGG GAT GTC TTA AAC TCT TTC AAT CCT AGA  
1380

GTT CCT GGA ATT AAG GGT GAA ATT CTT GGT GGC GAA ACA TGC TTA TGG AGT GAA ATG AAT  
1440

GAT GAT TCT ACT TAA TTC TAA AGA CTT TGG ACA AGA AAT AGT GCA TTT GCT GAA AGA CTT  
1500

TGG AAC ACT GAT GCT GCT AAC AAT GAA ACT TAC AAA ACT AGA GCT TTA GTT AGC AGA ATG  
1560

GTC TTT ATG CAA CAC CGT TTA ACT GCT AGA GGA ATC CCT GCT TCT CCT GTA ACA GTT GGT  
1620

ATT TGT GAA TAA AAC CTT TCT CTC TGC TAC AAT TGA  
1656

ttctaaatataaarattaaataaatattttaagaaatattttaagaatatttagtataaaaactgtattttaattga  
1735

taaaaaaaaaataaatattattattaattgaatttagctaaaaaaaaaaaaaaaaaaaaa  
1798

Figur 2: Die Aminosäure-Sequenz einer sauren Hydrolase aus dem Ciliaten *Tetrahymena*:

MQKILLITFL	LGIALAQITP	GVDPISAKVM	PKPKNYTYGD	LSLLVTDPCG	50
VSYPSPVGSG	KVPNHVYQII	GFYTLNIFNS	NENSCAMQR	EYKNETTIEK	100
MRRQLQHSQNI	VFDIFIQDAA	LATADTLEDE	YYDLQIYNTT	YWKLTANKYV	150
GLLRGLETYS	QLFTQDEDTE	DWYLNIPIS	IQDQPDYIYR	GLMIDSARHF	200
LSVETILKTI	DSMLFNKLN	LHWHITDES	FPFPLKSFPN	ITKYGAYSKK	250
KQYSFEDIQY	IVDQALNKG	QVIPEVDSPG	HAFSWARSPQ	FSSIGLLCDQ	300
YNGQLDPTLN	LYTAVKGIM	EDMNTQFYTA	KYVHFGGDEV	EEQCWNKRPE	350
IKEFMNQNNI	STYTDLQNY	RKNQVNIWKS	INATKPAIFW	ADSNTLKYGP	400
DDIIQWWGST	HDFSSIKDLP	NKIILSFYDN	TYLDVGEGNR	YGGSYGSMYN	450
WDVLNSFNPR	VPGIKGEILG	ETCLWSEMN	DDSTQFQRLW	TRNSAFAERL	500
WNTDAANNET	YKTRALVSRM	VFMQHRLTAR	GIPASPVTVG	ICEQNLSLCY	550
N					551

*Wey* 479  
*van* 90



# SEQUENZPROTOKOLL

<110> Hartmann, Marcus

<120> Saure Hydrolasen sowie diese kodierende DNA-Sequenz aus  
Ciliaten und deren Verwendung

<130> Hartmann, Marcus

<140>

<141>

<160> 3

<170> PatentIn Ver. 2.1

<210> 1

<211> 1656

<212> DNA

<213> Tetrahymena thermophila

<400> 1

```
atgcaaaaga tacttttaaat tacttttcctt cttggaatag ctctcgcctca aattactcct 60
ggcgttgacc ctatttcagc taagggttatg cctaaaccta agaattacac ttatggagat 120
ttgagcttac ttgtcactga tccttgcgga gtctcttaca gaccttctgt tgggctcagga 180
aaagtaccca accatgtcta tcaaattatt ggattctaca ctttgaatat tttcaattct 240
aacgaaaact cttgtgctat gtaaagagaa ttgtataaga atgaaacaac cattgaaaag 300
atgcgtagat tacaacattc ctaaaatata gtcttcgata tttttatcta agacgctgct 360
ttggccactg cagacacact cgaagacgaa tattatgatt tataaattta taataccaca 420
tattggaaat tgactgctaa caaatatgtt ggtttactcc gtgggtttaga aacttactct 480
caattattca cttaagacga agacactgaa gattggtatt tgaataacat ccctatttct 540
attcaagatt aaactgacta catctacaga ggtcctatga tagattcagc cagacatttc 600
tatcagttg aaactatttt aaaaactatt gattctatgt tattcaacaa gttgaatgtt 660
ctccattggc acatcactga tactgaatcc ttccccttcc ctcttaaate attcccta 720
attactaaat atggagccta ctctaagaag aaacaatata gcttcgaaga catttaatac 780
attgtagact aagctctcaa caagggtatt taagttattc ctgaagtcga ttctccagga 840
cacgcttttt catgggctag atctccttaa ttctctagta ttgggtctatt atgtgattaa 900
tataatggat agttagaccc aacactaaat ttaacttaca ctgctgttaa gggattatg 960
gaagatatga atacttaatt ctacactgct aagtatgttc attttggtgg tgatgaagtt 1020
gaagaataat gctggaataa acgccctgaa attaaggaat tcatgaatta aaataacatc 1080
tctacatata ctgatttgta gaattattac agaaagaact aagttaacat ttggaaatca 1140
attaatgcta ctaagcctgc tattttcttg gcagattcaa atactttgaa atatggctct 1200
gatgatatta ttcaatggtg gggatctact catgattttt cttcaatcaa agatcttctc 1260
aacaaaataa ttttatcttt ctatgataat acttattttg atgttggtga gggaaataga 1320
tatgggtgaa gttatggcag catgtataac tgggatgtct taaactcttt caatcctaga 1380
gttcctggaa ttaaggggtga aattcttggt ggcgaaacat gcttatggag tgaaatgaat 1440
gatgattcta cttaattcta aagactttgg acaagaaata gtgcatttgc tgaaagactt 1500
tggaacactg atgctgctaa caatgaaact tacaaaacta gagcttttagt tagcagaatg 1560
```

gtctttatgc aacaccgttt aactgctaga ggaatccctg cttctcctgt aacagttggt 1620  
 atttgtgaat aaaacctttc tctctgctac aattga 1656

<210> 2

<211> 549

<212> PRT

<213> Tetrahymena thermophila

<400> 2

Met Gln Lys Ile Leu Leu Ile Thr Phe Leu Leu Gly Ile Ala Leu Ala  
 1 5 10 15

Gln Ile Thr Pro Gly Val Asp Pro Ile Ser Ala Lys Val Met Pro Lys  
 20 25 30

Pro Lys Asn Tyr Thr Tyr Gly Asp Leu Ser Leu Leu Val Thr Asp Pro  
 35 40 45

Cys Gly Val Ser Tyr Arg Pro Ser Val Gly Ser Gly Lys Val Pro Asn  
 50 55 60

His Val Tyr Gln Ile Ile Gly Phe Tyr Thr Leu Asn Ile Phe Asn Ser  
 65 70 75 80

Asn Glu Asn Ser Cys Ala Met Gln Arg Glu Tyr Lys Asn Glu Thr Thr  
 85 90 95

Ile Glu Lys Met Arg Arg Leu Gln His Ser Gln Asn Ile Val Phe Asp  
 100 105 110

Ile Phe Ile Gln Asp Ala Ala Leu Ala Thr Ala Asp Thr Leu Glu Asp  
 115 120 125

Glu Tyr Tyr Asp Leu Gln Ile Tyr Asn Thr Thr Tyr Trp Lys Leu Thr  
 130 135 140

Ala Asn Lys Tyr Val Gly Leu Leu Arg Gly Leu Glu Thr Tyr Ser Gln  
 145 150 155 160

Leu Phe Thr Gln Asp Glu Asp Thr Glu Asp Trp Tyr Leu Asn Asn Ile  
 165 170 175

Pro Ile Ser Ile Gln Asp Gln Pro Asp Tyr Ile Tyr Arg Gly Leu Met  
 180 185 190

Ile Asp Ser Ala Arg His Phe Leu Ser Val Glu Thr Ile Leu Lys Thr  
 195 200 205

Ile Asp Ser Met Leu Phe Asn Lys Leu Asn Val Leu His Trp His Ile  
 210 215 220

Thr Asp Thr Glu Ser Phe Pro Phe Pro Leu Lys Ser Phe Pro Asn Ile  
 225 230 235 240

Thr Lys Tyr Gly Ala Tyr Ser Lys Lys Lys Gln Tyr Ser Phe Glu Asp  
 245 250 255

Ile Gln Tyr Ile Val Asp Gln Ala Leu Asn Lys Gly Ile Gln Val Ile  
 260 265 270

Pro Glu Val Asp Ser Pro Gly His Ala Phe Ser Trp Ala Arg Ser Pro  
 275 280 285

Gln Phe Ser Ser Ile Gly Leu Leu Cys Asp Gln Tyr Asn Gly Gln Leu  
 290 295 300

Asp Pro Thr Leu Asn Leu Thr Tyr Thr Ala Val Lys Gly Ile Met Glu  
 305 310 315 320

Asp Met Asn Thr Gln Phe Tyr Thr Ala Lys Tyr Val His Phe Gly Gly  
 325 330 335

Asp Glu Val Glu Glu Gln Cys Trp Asn Lys Arg Pro Glu Ile Lys Glu  
 340 345 350

Phe Met Asn Gln Asn Asn Ile Ser Thr Tyr Thr Asp Leu Gln Asn Tyr  
 355 360 365

Tyr Arg Lys Asn Gln Val Asn Ile Trp Lys Ser Ile Asn Ala Thr Lys  
 370 375 380

Pro Ala Ile Phe Trp Ala Asp Ser Asn Thr Leu Lys Tyr Gly Pro Asp  
 385 390 395 400

Asp Ile Ile Gln Trp Trp Gly Ser Thr His Asp Phe Ser Ser Ile Lys  
 405 410 415

Asp Leu Pro Asn Lys Ile Ile Leu Ser Phe Tyr Asp Asn Thr Tyr Leu  
 420 425 430

Asp Val Gly Glu Gly Asn Arg Tyr Gly Gly Ser Tyr Gly Ser Met Tyr  
 435 440 445

Asn Trp Asp Val Leu Asn Ser Phe Asn Pro Arg Val Pro Gly Ile Lys  
 450 455 460

Gly Glu Ile Leu Gly Glu Thr Cys Leu Trp Ser Glu Met Asn Asp Asp  
 465 470 475 480

Ser Thr Gln Phe Gln Arg Leu Trp Thr Arg Asn Ser Ala Phe Ala Glu  
 485 490 495

Arg Leu Trp Asn Thr Asp Ala Ala Asn Asn Glu Thr Tyr Lys Thr Arg  
 500 505 510

Ala Leu Val Ser Arg Met Val Phe Met Gln His Arg Leu Thr Ala Arg  
 515 520 525

Gly Ile Pro Ala Ser Pro Val Thr Val Gly Ile Cys Glu Gln Asn Leu  
 530 535 540

Ser Leu Cys Tyr Asn  
 545

<210> 3

<211> 1837

<212> DNA

<213> Tetrahymena thermophila

<400> 3

cagcagtaat aaaaaattct aaatatattg attgtagcta tgcaaaagat actttttaatt 60  
 acttttccttc ttggaatagc tctcgtctcaa attactcctg gcgttgaccc tatttcagct 120  
 aaggttatgc ctaaacctaa gaattacact tatggagatt tgagcttact tgtcactgat 180  
 ccttgcgag tctcttacag acctctgtt gggtcaggaa aagtacccaa ccatgtctat 240  
 caaattattg gatctctacac tttgaatatt ttcaattcta acgaaaactc ttgtgctatg 300  
 taaagagaat tgtataagaa tgaaacaacc attgaaaaga tgcgtagatt acaacattcc 360  
 taaaatatag tcttcgatat ttttatctaa gacgctgctt tggccactgc agacacactc 420  
 gaagacgaat attatgattt ataaatttat aataccacat attggaaatt gactgctaac 480  
 aaatatgttg gtttactccg tggtttagaa acttactctc aattattcac ttaagacgaa 540  
 gacactgaag attgggtattt gaataacatc cctattttcta ttcaagatta acctgactac 600  
 atctacagag gtcttatgat agattcagcc agacatttct tatcagttga aactatttta 660  
 aaaactattg attctatgtt attcaacaag ttgaatgttc tccattggca catcactgat 720  
 actgaatcct tccccctccc tcttaaatca ttccttaata ttactaaata tggagcctac 780  
 tctaagaaga aacaatacag cttcgaagac atttaataca ttgtagacta agctctcaac 840  
 aagggtattt aagttattcc tgaagtogat tctccaggac acgctttttc atgggctaga 900  
 tctccttaat tctctagtat tggctctatta tgtgattaat ataattggata gttagaccca 960  
 acactaaatt taacttacac tgctgttaag ggtattatgg aagatatgaa tacttaattc 1020  
 tacactgcta agtatgttca ttttggtggt gatgaagttg aagaataatg ctggaataaa 1080  
 cgccctgaaa ttaaggaatt catgaattaa aataacatct ctacatatat tgatttgtag 1140  
 aattattaca gaaagaacta agttaacatt tggaaatcaa ttaatgctac taagcctgct 1200  
 attttctggg cagattcaaa tactttgaaa tatggtcctg atgatattat tcaatggtgg 1260

ggatctactc atgatttttc ttcaatcaaa gatcttccta acaaaataat tttatctttc 1320  
tatgataata cttatttgga tgttggtgag ggaaatagat atggtggaag ttatggcagc 1380  
atgtataact gggatgtctt aaactctttc aatcctagag ttcctggaat taagggtgaa 1440  
attcttggtg gcgaaacatg cttatggagt gaaatgaatg atgattctac ttaattctaa 1500  
agactttgga caagaaatag tgcatttgct gaaagacttt ggaacactga tgctgctaac 1560  
aatgaaactt acaaaactag agcttttagtt agcagaatgg tctttatgca acaccgttta 1620  
actgctagag gaatccctgc ttctcctgta acagttggta tttgtgaata aaacctttct 1680  
ctctgctaca attgattcta aatataaara ttaaataaat attttaagaa atatttttaa 1740  
gaatatttta gtataaaaac tgtattttta ttgataaaaa aaatataaat attattatta 1800  
attgaatttt agctaaaaaa aaaaaaaaaa aaaaaaa 1837